Iranian Society of Animal Science

Journal of Livestock Science and Technologies

Shahid Bahonar University of Kerman

ISSN: 2322-3553 (Print) ISSN: 2322-374X (Online)

Paper type: Original Research

Modeling a latent variable for body size using morphometric traits in cattle

Morteza Mokhtari[*, 1], *Morteza Bitaraf Sani*[2], *Rouhollah Mirmahmoudi*[1], *Yadollah Badakhshan*[1], *Mozhgan Mazhari*[1]

[1] *Department of Animal Science, Faculty of Agriculture, University of Jiroft, P.O. Box 364, Jiroft, Iran*
[2] *Animal Science Research Department, Yazd Agricultural and Natural Resources Research and Education Center, Agricultural Research, Education & Extension Organization (AREEO), Yazd, Iran*

[*]Corresponding author,
E-mail address:
msmokhtari@ujiroft.ac.ir

Received: 09 Dec 2025,
Received in revised form: 07 Feb 2026,
Accepted: 24 Feb 2026,
Published online: 27 Feb 2026,

ORCID
Morteza Mokhtari
0000-0002-3971-628X
Morteza Bitaraf Sani
0000-0001-5694-0569
Rouhollah Mirmahmoudi
0000-0002-0022-0105
Yadollah Badakhshan
0000-0001-9738-225X
Mozhgan Mazhari
0000-0001-7719-2707

Abstract This study aimed to quantify a latent variable for body size (BS) in cattle by using six morphometric traits, including body height at withers (HW), body length (BL), hip width (HpW), chest depth (CD), shoulder width (SW), and chest width (CW). The statistical measures for goodness of fit, including comparative fit index (CFI), Tucker-Lewis Index (TLI), root mean square error of approximation (RMSEA), and standardized root mean square residual (SRMR), were 0.94, 0.91, 0.05, and 0.05, respectively, and appropriately indicate the adequacy of the confirmatory factor model proposed for the latent variable of BS. The standardized factor loadings of HW, BL, HpW, CD, SW, and CW for describing BS were 0.83, 0.76, 0.82, 0.89, 0.80, and 0.40, respectively, and statistically significant (P<0.01), implying that the observed variables were appropriate indicators of the corresponding BS latent trait. All correlations among morphometric traits were positive and statistically significant (P<0.01), ranging from 0.26 (CW-SW) to 0.77 (HW-CD). The correlations between the BS latent trait and the considered morphometric traits were also positive and statistically significant (P<0.01); ranged from 0.42 (CW-BS) to 0.93 (CD-BS). It was concluded that the proposed confirmatory factor analysis model showed adequate fit indices for constructing the BS latent trait, implying that the proposed framework adequately captures the underlying relationships among the observed variables. Overall, the study provided a robust framework for applying BS in contexts such as phenotypic evaluations, where a latent construct can capture the complexity of morphometric traits more effectively than individual traits alone.

Keywords: cattle, factor analysis, latent variable, morphometric traits

Introduction

Multivariate models play a crucial role in quantitative genetics. Multivariate models simultaneously analyze several traits, leveraging genetic covariances between traits to improve the accuracy of estimates of genetic parameters and breeding values. These models also provide a deeper understanding of the genetic background underlying complex, interrelated phenotypes, which is crucial for both theoretical advancements and practical applications in animal genetics (Gianola and Sorensen, 2004). The value of a selection candidate as a potential parent for the next generation is often evaluated based on multiple traits, such as growth, body conformation, and production (Gianola and Sorensen, 2004). A key characteristic of multivariate models is that the number of parameters increases with the number of traits included; this parameter inflation can compromise the feasibility of genetic analyses (Silva et al., 2021). Parameter inflation can bias estimates of genetic parameters and alter the perceived genetic potential of traits, thereby affecting selection decisions (Lynch and Walsh, 1998).

Appropriate statistical techniques, such as structural equation modeling (SEM), were used to address such concerns. The SEM, introduced by Wright (1921), is an advanced multivariate statistical technique that includes a variety of models to represent the relationships among observed variables and also allows for modeling relationships between the latent and observed variables, enabling quantitative evaluation of theoretical or measurement models (Schumacker and Lomax, 1996). Statistically, latent variables are constructs that cannot be directly measured but can be inferred from multiple observed phenotypes (Schumacker and Lomax, 1996). Latent variable modeling (LVM) is one of the primary applications of SEM across numerous scientific disciplines (Schumacker and Lomax, 1996). The LVM, as a dimension-reduction technique, can be applied to phenotypic data to reduce dataset dimensionality and computational complexity in such a way that many phenotypes are combined to represent a few underlying concepts of interest (Leal-Gutierrez et al., 2018). The latent traits are assumed to be unobserved, but they are believed to explain the covariation among the observed variables.

Silva et al. (2021) showed that a set of latent variables can effectively capture the underlying structure of phenotypic data, thereby mitigating the complexity caused by model over-parameterization. The LVM offers a powerful approach for investigating biologically complex phenomena by reducing data dimensionality, allowing many phenotypes to be combined into a smaller set of underlying concepts of interest (Leal-Gutierrez et al., 2018). Confirmatory factor analysis (CFA) is one of the methods within SEM that can be applied to reduce the dimensionality of phenotypic data in genomic studies (Yu et al., 2019). The CFA is typically employed to test the hypothesis that a set of observed variables is associated with a smaller set of latent factors, also referred to as latent variables or latent traits (Momen et al., 2021).

Morphometric traits are important in the phenotypic characterization of any breed or population and are considered to evaluate the body conformation of different livestock breeds, compare individual growth patterns, and describe individuals or populations more accurately (Yakubu et al., 2021). These studies are often regarded as the first step in conserving animal genetic resources (Bousbia et al., 2021). Morphometric traits in cattle, such as stature, chest width, body depth, and rump angle, are often highly correlated, which can lead to collinearity problems in genetic and phenotypic analyses, which can be mitigated by introducing a latent variable through factor analysis, which summarizes the shared variance among these traits into a single composite measure (Xu et al., 2022). The present study aimed to model and identify the latent trait of body size (BS) using six morphometric traits in Baoule dairy cattle. Baoule cattle (*Bos taurus*), locally called Lobi cattle, are an important taurine population located in the Southwest of Burkina Faso, a region known to be tsetse-challenged, and are primarily raised for both dairy and beef production (Yougbare et al., 2021).

Materials and methods

Data source and traits

The data used in the present study were obtained from the morphometric records provided by Yougbare et al. (2021) in purebred and crossbred Baoule cattle of Burkina Faso. In the present study, the phenotypic records of six morphometric traits on 223 Baoule cattle, including body height at withers (HW), body length (BL), hip width (HpW), chest depth (CD), shoulder width (SW), and chest width (CW), were used. Descriptive statistics for morphometric traits are presented in Table 1. Among the animals, 33.82% were male and 66.18% were female. The proportions of 1-year, 2-year, 3-year, 4-year, 5-year, 6-year, and 7-year-old animals were 20.59%, 13.72%, 11.76%, 15.69%, 12.25%, 9.90%, and 16.18%, respectively. Additional details regarding the animals and phenotypes have been reported by Yougbare et al. (2021).

Table 1. Descriptive statistics for the studied morphometric traits in Baoule cattle

Trait [1]	Mean	S.D.	C.V. (%)	Minimum	Maximum
HW (cm)	98.47	9.79	9.94	77.00	124.00
BL (cm)	79.52	9.99	12.56	30.00	108.00
HpW (cm)	29.52	3.85	13.04	13.00	39.00
CD (cm)	50.13	4.54	9.06	36.00	61.00
SW (cm)	25.54	3.51	13.74	16.00	35.00
CW (cm)	12.28	2.25	18.32	7.00	27.00

[1] HW: height at withers, BL: body length, HpW: hip width, CD: chest depth, SW: shoulder width, CW: chest width.

Statistical analysis

Latent variable modeling by confirmatory factor analysis
The theoretical model considered for constructing the BS latent variable from six morphometric traits is presented in Figure 1. The CFA was employed to identify the underlying latent factors contributing to variation in the studied morphometric traits. To perform a CFA aimed at extracting a latent factor related to BS, it is required to establish a measurement model that defines the relationships between the observed variables and the latent factor(s). The CFA model outlined below was considered in this study.

$$\mathbf{X} = \boldsymbol{\Lambda}\boldsymbol{\xi} + \boldsymbol{\delta}$$

Here, **X** is the matrix of morphometric traits (HW, BL, HpW, CD, SW, and CW), while **ξ** is the vector of the latent factors. The **Λ** matrix contains the factor loadings that link these latent factors to the morphometric traits, and **δ** is the residual vector. Residuals were assumed to be uncorrelated. Normality of the morphometric characteristics was evaluated by the Kolmogorov-Smirnov test (SAS, 2010). The CFA model was fitted using maximum likelihood estimation with the *lavaan* R package (Rosseel, 2012) (R Development Core Team, 2025). The goodness of fit for the hypothesized model for constructing BS was evaluated using four statistical measures, including the root mean square error of approximation (RMSEA), standardized root mean square residual (SRMR), Tucker–Lewis Index (TLI), and comparative fit index (CFI) (Bentler, 1990). The closer CFI and TLI are to one, and the closer SRMR is to zero, the more favorable the model. The model was tested for bias through permutation, using the bootstrapLavaan function of the lavaan package (Rosseel, 2012) with 5,000 bootstrap draws. The estimated values for the BS latent variable obtained from the CFA followed a standard normal distribution with a mean of zero and a standard deviation of one. Pearson's correlations between the morphometric traits and BS were also computed (SAS, 2010).

Results and discussion

The BS latent trait and estimated factor loadings

Figure 1 presents a comprehensive representation of the hypothesized model for BS, illustrating the underlying relationships within the system under investigation and incorporating BS as a latent variable at the center of the six measured morphometric variables. The goodness-of-fit indices, including CFI, TLI, RMSEA, and SRMR, were 0.94, 0.91, 0.05, and 0.05, respectively. The goodness-of-fit measures clearly demonstrate the adequacy of the CFA proposed for the latent variable of BS. The primary aim of using CFA is to reduce dimensionality in a multivariate context, making the interpretation of interrelationships among variables simpler when several phenotypes are simultaneously involved (Silva et al., 2021). Morphometric traits have also been used for constructing a latent variable of body size in pigs (Sanjari Banestani et al., 2023) and in beef heifers (Anas et al., 2025).

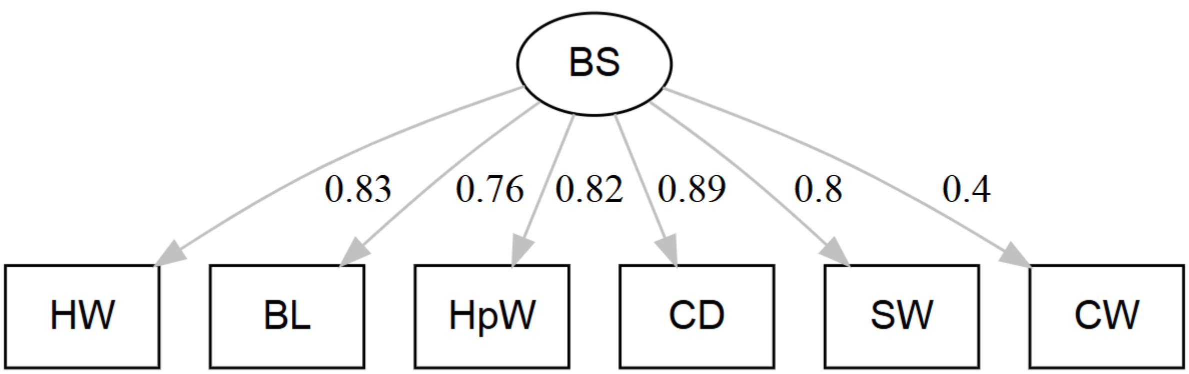

Figure 1. Representation of the latent variable of body size (BS) and the respective relationships with height at withers (HW), body length (BL), hip width (HpW), chest depth (CD), shoulder width (SW), and chest width (CW). For each observed variable, the standardized estimation of loadings is presented.

As shown in Table 2, the standardized factor loadings for the observed variables of HW, BL, HpW, CD, SW, and CW were 0.83, 0.76, 0.82, 0.89, 0.80, and 0.40, respectively, and significant (P<0.01), implying that the observed variables were appropriate indicators of the corresponding BS latent variable. The standardized factor loadings can be regarded as regression coefficients, representing the amount of association between measured traits and the underlying latent variable (Penagaricano et al., 2015).

In the present study, the factor loadings further substantiated the validity of the measurement model. Except for CW, which showed a comparatively weaker loading (0.40), all indicators demonstrated strong and statistically significant associations with BS. This pattern underscores the robustness of HW, BL, HpW, CD, and SW as primary contributors to BS, while CW, though significant, appears to play a more limited role in defining the latent construct. Such variation in indicator strength highlights the multidimensional nature of BS and suggests that certain traits may exert greater influence in shaping the overall construct. By applying five morphometric traits under the CFA, Sanjari Banestani et al. (2023) constructed a latent variable of body size in Yorkshire pigs and reported the factor loadings for the observed variables of body length, body height, chest width, chest girth, and tube girth as 0.62, 0.84, 0.57, 0.54, and 0.63, respectively. By applying CFA, Anas et al. (2025) constructed a latent variable of body size in beef heifers and estimated loadings of 0.699, 0.901, 0.828, 0.799, 0.704, 0.727, and 0.911 for body length, body weight, flank girth, heart girth, hip height, hip width, and mid girth, respectively. Although there were species differences, comparison of the loading factors reported for morphometric traits that describe body size in other livestock species is of interest. In the present study, the estimated factor loading for HW (0.83) was similar to that reported by Sanjari Banestani et al. (2023) for body height in Yorkshire pigs (0.84). The corresponding estimated factor loading for BL (0.76) in our study was higher than those reported by Sanjari Banestani et al. (2023) in Yorkshire pigs (0.62) and by Anas et al. (2025) in beef heifers (0.699). Differences in factor loadings across studies may be due to species, breed, age, or management conditions.

Table 2. Standardized factor loadings of morphometric traits used for describing the BS latent variable

Trait [1]	Factor loadings ± standard error	99% Confidence interval
HW	0.83±0.03	0.78-0.88
BL	0.76±0.05	0.66-0.87
HpW	0.82±0.04	0.74-0.89
CD	0.89±0.05	0.77-0.98
SW	0.80±0.03	0.74-0.86
CW	0.40±0.09	0.22-0.58

[1] HW: height at withers, BL: body length, HpW: hip width, CD: chest depth, SW: shoulder width, CW: chest width.

The loading factors for BS were positive and consistent with the theoretical model, in which the measured morphometric traits are expected to be positively related to BS. All factor loadings were positive, indicating that an increase in any measured variable is associated with an increase in the latent variable BS. Among the measured traits, CD exhibited the strongest relationship with BS based on the estimated loadings. The CD refers to the vertical distance from the top of the withers to the bottom of the chest, at the point where the sternum is located, and is often considered an indicator of the overall body volume and mass. A deeper chest typically correlates with greater body size.

A positive factor loading confirms that an increase in a measurable variable is accompanied by an increase in BS. This positive direction was anticipated, as the BS construct primarily reflects body measurement traits (Kominakis et al., 2017). Overall, the results suggested that the latent variable BS might be important in explaining the interrelationships among the measured morphometric traits. These findings highlight the importance of incorporating latent variables when modeling complex systems.

Correlations among the traits

All phenotypic correlation coefficients between body dimensions and BS (Table 3) were positive and significant (P<0.01), ranging from 0.26 (CW-SW) to 0.77 (HW-CD). The correlations between the BS latent trait and body dimensions were also positive and significant (P<0.01) and ranged from 0.42 (CW-BS) to 0.93 (CD-BS). The findings showed that morphometric traits in cattle are strongly and positively correlated, and that the BS latent trait captures this shared variance even more strongly. This confirms the presence of collinearity among morphometric traits and highlights the usefulness of a latent variable approach for summarizing them.

Table 3. Pearson's correlations among the studied traits

Trait [1]	HW	BL	HpW	CD	SW	CW	BS
HW	-	0.66 **	0.65 **	0.77 **	0.61 **	0.39 **	0.87 **
BL		-	0.65 **	0.64 **	0.60 **	0.31 **	0.80 **
HpW			-	0.71 **	0.73 **	0.18 **	0.85 **
CD				-	0.70 **	0.43 **	0.93 **
SW					-	0.26 **	0.83 **
CW						-	0.42 **
BS							-

[1] HW: height at withers, BL: body length, HpW: hip width, CD: chest depth, SW: shoulder width, CW: chest width, BS: body size.
** Statistically significant at P<0.01.

The latent BS trait showed even stronger correlations with morphometric traits (0.42–0.93). The highest correlation (CD-BS = 0.93) suggested that chest depth is a particularly strong indicator of BS. The lowest correlation (CW-BS = 0.42) still indicated a moderate relationship, meaning chest width contributes but less strongly than other traits. These results validated the latent BS trait as a robust summary measure that integrates multiple correlated dimensions. Yakubu et al. (2021) reported that the strong positive correlations among morphometric traits in cattle may be valuable for selection purposes, given the similar gene action of positively correlated traits. They also pointed out that this approach can be utilized for the genetic improvement of indigenous stock, particularly in rural areas where resources for large-scale breeding programs are limited.

Conclusion

The present results affirmed the appropriateness of modeling BS as a latent variable and suggest that the considered morphometric traits collectively provide a coherent representation of the construct. From a statistical perspective, these findings suggest that BS might be considered as an integrative measure to evaluate morphometric traits in a theoretical framework. The consistently positive and significant correlations among the morphometric traits highlighted their biological coherence, reinforcing the suitability of BS as a composite trait rather than relying on single measurements. It should be noted that BS is treated as a latent construct without validation. Therefore, subjective body size scores provided by trained classifiers or objective reference measures such as body weight are required to support any possible application of the body size latent variable. Overall, the study provided a robust statistical framework for describing BS in contexts such as phenotypic evaluations, where a latent construct can capture the complexity of morphometric traits more effectively than individual traits alone.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgements

We would like to acknowledge the University of Jiroft for financial support of the study.

References

Anas, M., Zhao, B., Yu, H., Dahlen, C.R., Swanson, K.C., Ringwall, K.A., Hulsman Hanna, L.L., 2025. Multi-trait phenotypic modeling through factor analysis and Bayesian network learning to develop latent reproductive, body conformational, and carcass-associated traits in admixed beef heifers. *Frontiers in Genetics* 16, 1551967.

Bentler, P.M., 1990. Comparative fit indexes in structural models. *Psychological Bulletin* 107, 238-246.

Bousbia, A., Boudalia, S., Gueroui, Y., Hadded, K., Bouzaoui, A., Kiboub, D., Symeon, G., 2021. Use of multivariate analysis as a tool in the morphological characterization of the main indigenous bovine ecotypes in northeastern Algeria. *PLoS One* 16(7), e0255153.

Gianola, D., Sorensen, D., 2004. Quantitative genetic models for describing simultaneous and recursive relationships between phenotypes. *Genetics* 167, 1407-1424.

Kominakis, A., Hager-Theodorides, A.L., Zoidis, E., Saridaki, A., Antonakos, G., Tsiamis, G., 2017. Combined GWAS and guilt-by-association-based prioritization analysis identifies functional candidate genes for body size in sheep. *Genetics Selection Evolution* 49, 41.

Leal-Gutierrez, J.D., Rezende, F.M., Elzo, M.A., Johnson, D., Penagaricano, F., Mateescu, R.G., 2018. Structural equation modeling and whole-genome scans uncover chromosome regions and enriched pathways for carcass and meat quality in beef. *Frontiers in Genetics* 9, 1-13.

Lynch, M., Walsh, B., 1998. Genetics and Analysis of Quantitative Traits. Oxford University Press, Sinauer Associates, Inc., Sunderland, MA.

Momen, M., Bhatta, M., Hussain, W., Yu, H., Morota, G., 2021. Modeling multiple phenotypes in wheat using data-driven genomic exploratory factor analysis and Bayesian network learning. *Plant Direct* 5, e00304.

Penagaricano, F., Valente, B. D., Steibel, J. P., Bates, R. O., Ernst, C. W., Khatib, H., Rosa, G. J. M., 2015. Searching for causal networks involving latent variables in complex traits: Application to growth, carcass, and meat quality traits in pigs. *Journal of Animal Science* 93, 4617-4623.

R Development Core Team, 2025. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rosseel, Y., 2012. lavaan: An R package for structural equation modeling. *Journal of Statistical Software* 48, 1-36.

Sanjari Banestani, E., Esmailizadeh, A., Momen, M., Ayatollahi Mehrgardi, A., Mokhtari, M., 2023. Genome-wide association study identifies significant SNP and related genes associated with body size in Yorkshire pigs using latent variable modelling. *Journal of Agricultural Science* 161, 599-605.

SAS. 2010. SAS User's Guide, Version 9.4. Statistical Analysis System. SAS Institute Inc. Cary, North Carolina, USA.

Schumacker, E.R., Lomax, G.R., 1996. A Beginner's Guide to Structural Equation Modeling. Erlbaum, Mahwah, NJ.

Silva, H.T., Paiva, J.T., Botelho, M.E., Carrara, E.R., Lopes, P.S., Silva, F.F., Veroneze, R., Sterman Ferraz, J.B., Eler, J.P., Mattos, E.C., Gaya, L.G., 2021. Searching for causal relationships among latent variables concerning performance, carcass, and meat quality traits in broilers. *Journal of Animal Breeding and Genetics* 139, 181-192.

Wright, S., 1921. Correlation and causation. *Journal of Agricultural Research* 20, 557-585.

Xu, L., Luo, H., Zhang, X., Lu, H., Zhang, M., Ge, J., Zhang, T., Yan, M., Tan, X., Huang, X., et al., 2022. Factor analysis of genetic parameters for body conformation traits in dual-purpose Simmental cattle. *Animals* 12, 2433.

Yakubu, A., Hingir, A.V., Abdullah, A-R., 2018. Multivariate analysis of sexual dimorphism in the morphometric traits of Muturu cattle in north central Nigeria. *Nigerian Journal of Genetics* 32, 8-15.

Yougbare, B., Soudre, A., Ouedraogo, D., Zoma, B.L., Tapsoba, A.S.R., Sanou, M., et al., 2021. Genome-wide association study of trypanosome prevalence and morphometric traits in purebred and crossbred Baoule cattle of Burkina Faso. *PLoS One* 16(8), e0255089.

Yu, H., Campbell, M.T., Zhang, Q., Walia, H., Morota, G., 2019. Genomic Bayesian confirmatory factor analysis and Bayesian network to characterize a wide spectrum of rice phenotypes. *Genes, Genomes Genetics* 9, 1975-1986.